#24

Docket No.: 042390.P5700

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE
BEFORE THE BOARD OF PATENT APPEALS AND INTERFERENCES

In re Application of: Jiang et al.

Assignee: Intel Corporation

Application No.: 09/470,741

Filed: December 20, 1999

For: METHOD AND APPARATUS FOR PERFORMING VIDEO IMAGE DECODING

Examiner: J. Wu

Art Group: 2623

APPEAL BRIEF
IN SUPPORT OF APPELLANTS' APPEAL
TO THE BOARD OF PATENT APPEALS AND INTERFERENCES

Mail Stop Appeal Brief - Patents
Commissioner for Patents
P.O. Box 1450
Alexandria, VA 22313-1450

Sir:

Appellants hereby submit this Appeal Brief, in triplicate, in support of Appellants' Appeal from final rejection of the pending claims in the above-captioned case.

A Notice of Appeal is being filed together with this Appeal Brief.

The fees set forth in 37 CFR § 1.17(b) and (c) accompany this Appeal Brief.

An oral hearing is NOT desired.

Appellants respectfully request consideration of this Appeal by the Honorable Board of Patent Appeals and Interferences, and allowance of the claims of the subject application.

Please charge any fees and/or credit any overcharges to Deposit Account No. 02-2666.

I. REAL PARTY IN INTEREST

The invention is assigned to Intel Corporation of 2200 Mission College Boulevard, Santa Clara, California 95052

II. RELATED APPEALS AND INTERFERENCES

To the best of Appellants' knowledge, there are no appeals or interferences related to the present appeal that will directly affect, be directly affected by, or have a bearing on the Board's decision in the subject appeal.

III. STATUS OF THE CLAIMS

Claims 1-9, 12-21, and 23-34 are currently pending in the subject application. These claims were finally rejected in the Final Office Action mailed March 25, 2003, and are the subject of this appeal. The Examiner confirmed the final rejection of these claims in an Advisory Action mailed May 12, 2003.

In the Final Office Action, the Examiner raised four grounds of rejection. First, the Examiner has rejected claims 1-7, 9, 11-12, 16-19, 21-24, 28-30, and 32-34 under 35 USC § 103 as being obvious in view of Vetro et al. ("Frequency Domain Down-Conversion of HDTV Using an Optimal Motion Compensation Scheme," Journal of Imaging Systems and Technology, Vol. 9, No. 4, August 1998, pp. 274-282) in view of Ng (U.S. Patent No. 5,262,854) and Bose et al. (U.S. Patent No. 6,215,822). Second, the Examiner has rejected claims 14 and 26 under 35 USC § 103 as being rendered obvious by the combination of Vetro et al., Bose et al., Ng, and Dugad et al. ("A Fast Scheme for Altering Resolution in the Compressed Domain," IEEE Computer Science Conference on Computer Vision and Pattern Recognition, June 1999, pp. 213 - 218). Third, the Examiner has rejected claims 8 and 31 under 35 USC § 103 as being rendered obvious

by the combination of Vetro et al. and Ng, in further view of Kim et al. (U.S. Patent No. 6,175,592). Fourth, the Examiner has rejected claims 15 and 27 under 35 USC § 103 as being obvious over the combination of Vetro et al., Ng, Bose et al., and Dugad et al., in further view of Rosman et al. (U.S. Patent No. 6,222,550). Appellants respectfully traverse each of these grounds of rejection.

IV. STATUS OF AMENDMENTS

In response to the Final Office Action, Appellants filed on April 16, 2003 an Amendment, After Final Rejection, Under Rule 116 (hereinafter, "Amendment"). In the Advisory Action, the Examiner indicated that the Amendment would be entered by the Examiner.

In response to the Final Office Action and the Advisory Action, a Notice of Appeal is being timely filed together with this Appeal Brief. A copy of all of the claims on appeal is attached hereto as Appendix A.

V. SUMMARY OF THE INVENTION

Conventionally, a video decode and display system normally designed for a given maximum resolution will typically not operate on bit streams that specify higher video resolutions. Likewise, if a user chooses to view the video in a smaller window, downscaling the bit stream is conventionally achieved at display time and, therefore, full resolution decoding still occurs. Since full resolution decoding followed by downscaling adds cost in the form of additional computation, additional memory, additional memory bandwidth, and complex downscaling at display time, it would be desirable if downscaling of a bit stream could be accomplished without full resolution decoding. (Specification, page 6, lines 5-11).

FIG. 7 of the subject application is a schematic diagram illustrating an embodiment of a hardware motion compensation engine that may be employed to implement an embodiment of a method of performing video decoding in accordance with the invention. For example, a three-dimensional (3D) pipeline may be employed to efficiently perform motion compensation, as illustrated in FIG. 8 of the subject application, although other hardware platforms other than a 3D pipeline may be employed to implement embodiments of a method of performing video decoding in accordance with an embodiment of the invention. FIGs. 7 and 8 of the subject application are described in more detail hereinafter. (Specification, page 6, line 12 to page 7, line 1; Figures 7 and 8).

FIG. 1 of the subject application is a block diagram illustrating an embodiment of conventional pipeline for performing video image decoding. FIG. 2 of the subject application is a block diagram illustrating an embodiment of an apparatus for performing video decoding in accordance with the present invention. In one embodiment in accordance with the invention, as shall be described in greater detail hereinafter, a compressed video image in the frequency domain is downsampled at 230 and then inverse transformed at 220. Motion compensation is performed on the downsampled image in the spatial domain at 210. Alternatively, as shall also be described in greater detail hereinafter, the compressed image in the frequency domain may be inverse transformed at 240 and then downsampled in the spatial domain and motion compensated. Although the invention is not limited in scope in this respect, one example of a compressed video image in the frequency domain comprises a discrete cosine transform (DCT) image. Likewise, although the invention is not limited in scope in this respect, such a DCT image may comply with the MPEG2 specification, as shall be described in greater detail hereinafter. In this context, although MPEG2 is referred to, including the aspect that divides an image into 16 x16 macroblocks, the invention is not limited in scope to employing MPEG,

including MPEG2, to employing macroblocks of this particular size and shape or even to employing macroblocks at all. (Specification, page 7, lines 2-16; Figures 1 and 2).

As illustrated in FIG. 2 of the subject application, and as shall be described in greater detail, the DCT image may be downsampled before being delivered to the motion compensation engine. Likewise, as indicated above, downsampling may be applied in this embodiment either before the inverse DCT, such as at 230, or after the inverse DCT, such as at 240, depending upon a variety of factors. In this particular embodiment, although the invention is not limited in scope in this respect, the blocks illustrated in FIG. 2 of the subject application prior to the vertical line are implemented in software and the blocks after the vertical line are implement in hardware. Conventionally, such video processing to accomplish downsampling would be performed in hardware; however, an advantage of an embodiment of a method of performing video decoding in accordance with the present invention is that it provides the capability to perform the processing in software due at least in part to greater processing efficiency in comparison with conventional approaches. Therefore, one advantage of this approach is that it provides greater flexibility. In such an embodiment, the decoder software may transfer the downsampled prediction error to the motion compensation hardware and the motion vectors may be adjusted substantially in accordance with the downsampling ratio, as explained hereinafter. In this particular embodiment, although, again the invention is not limited in scope in this respect, downsampling ratios of 1:1, 2:1, 4:1, and 8:1, along either of the horizontal, vertical or both directions, may be supported. In this particular embodiment, where MPEG2 is employed, the downsampling ratio is limited to no more than 8:1 due to the native eight-by-eight MPEG2 block size. However, this limitation may not apply in alternative embodiments. Furthermore, in alternative embodiments, even for MPEG2, downsampling ratios other than a power of two may

be implemented, such as, for example, 3:1. (Specification, page 7, line 17 to page 8, line 18; Figure 2).

As illustrated in FIG. 2 of the subject application, the motion compensation hardware may operate directly on the downsampled bit stream. In this particular embodiment, where MPEG2 is employed, the downsampling ratio may be n, where n equals 1, 2, 4, and 8. In a motion compensation process, a motion vector of a processed macroblock specifies the relative distance of reference data from the processed macroblock. Let (V_x, V_y) = (vector[r][0], vector[r][1]) be the original motion vector for a macroblock, where V_x and V_y, the horizontal and vertical components of the motion vector, are in the form of 16-bit signed value, although the invention is not limited in scope in this respect. According to the MPEG2 standard, the least significant bit (LSB) of Vx and Vy is used to indicate the half-pixel resolution reference. Denote the whole pixel motion displacement for the luminance component by $\left(D_x^Y, D_y^Y\right)$, and the fractional offset for the luminance component by $\left(F_x^Y, F_y^Y\right)$. Due to limited subpixel precision, the fractional offset $\left(F_x^Y, F_y^Y\right)$ may also be called as the half-pixel offset flag. When there is no downsampling to the bit stream, these may be calculated from the motion vector as follows:

$$\begin{cases} D_x^Y = V_x >> 1, \\ D_y^Y = V_y >> 1, \end{cases} \qquad [1]$$

and

$$\begin{cases} F_x^Y = V_x \,\&\, 1, \\ F_y^Y = V_y \,\&\, 1. \end{cases} \qquad [2]$$

where ">>" indicates a right shift operation and "&" indicates a "logic AND" operation. If F_x^Y is set, or precisely is non-zero, horizontal interpolation, such as computing an average, may be applied to the reference pixels. If F_y^Y is set, vertical interpolation, such as computing an average, may be applied to the reference pixels. If both are set, interpolations along both directions may be applied. (Specification, page 9, line 1 to page 10, line 3; Figure 2).

The chrominance motion displacement may also derived from the same set of motion vector signal information. For YUV 4:2:0 color space format, for example since the dimension of chrominance (Cb, Cr) pictures is half of that of the luminance component picture along both horizontal and vertical directions, the whole pixel displacement $\left(D_x^C, D_y^C\right)$ and fractional offset $\left(F_x^C, F_y^C\right)$ for the chrominance components of the processed macroblock may be determined as follows:

$$\begin{cases} D_x^C = \left(V_x / 2\right) >> 1, \\ D_y^C = \left(V_y / 2\right) >> 1, \end{cases} \qquad [3]$$

and

$$\begin{cases} F_x^C = \left(V_x / 2\right) \& 1, \\ F_y^C = \left(V_y / 2\right) \& 1, \end{cases} \qquad [4]$$

where symbol '/' denotes regular integer division with truncation of the result toward zero. Notice that in this example the chrominance fractional offset is also in half chrominance pixel resolution. (Specification, page 10, lines 4-15).

Ignoring the luminance and chrominance superscripts of terms $\left(D_x, D_y\right)$ and $\left(F_x, F_y\right)$, the

motion prediction operation may be, in one embodiment, implemented with simple adders and shifters as the following pseudo-code illustrates:

```
MC_Prediction(p, q)
{
    if( Fx !=1 && Fy != 1)                    /* full-pel prediction in both
    directions */
        P[q][p] = R[n][m];
    if( Fx ==1 && Fy != 1)                    /* full-pel vertical, half-pel
    horizontal */
        P[q][p] = (R[n][m] + R[n][m+1] ) // 2;
    if( Fx !=1 && Fy == 1)                    /* half-pel vertical, full pel
    horizontal */
        P[q][p] = (R[n][m] + R[n+1][m] ) // 2;
    if( Fx ==1 && Fy == 1)                    /* half-pel prediction in both directions */
        P[q][p] = (R[n][m] + R[n][m+1] + R[n+1][m] + R[n+1][m+1]) // 4;
} // end Mc_Prediction(p, q)
```

In this example, the division symbol "//" denotes rounding up to the next larger integer (rounding away from zero). Symbols p and q represent integer indices in the destination image along horizontal and vertical directions, respectively. Symbols m and n represent integer indices in the reference image along horizontal and vertical directions, respectively. The reference pixel location (m, n) may be derived from the motion vector displacement term (D_x, D_y).

(Specification, page 10, line 16 to page 11, line 16).

In this particular embodiment of the invention, the motion-compensated prediction applied to the downsampled bit stream is performed directly using the downsampled reference images and the original motion vectors decoded from the compressed bit stream. The motion vectors used in the prediction may also be specified by the motion displacement (D_x^Y, D_y^Y), (D_x^C, D_y^C) and motion fractional offset (F_x^Y, F_y^Y), (F_x^C, F_y^C) with reference to the downsampled image. Contrary to conventional motion fractional offset that is only a one bit value in MPEG2, as previously described, more precision is preserved for (F_x^Y, F_y^Y), (F_x^C, F_y^C) in a downsampling operation in this particular embodiment in accordance with the invention. Consequently, the simple averaging operation described above may be replaced by more accurate interpolation operations. In one embodiment, for example, a bilinear interpolation unit may be used in the motion prediction calculation of motion compensation, although the invention is not limited in scope in this respect. The video or image reconstruction quality may also be improved by using a higher order interpolation unit. A bilinear interpolator typically employs more hardware than an averaging based interpolator. However, it is a common feature that may be provided as part of a state-of-the-art graphics controller hardware. For example, it may be found in the texture pipeline of a three-dimensional (3D) rendering engine or an image processor for image scaling or filtering. In one embodiment, therefore, as illustrated in FIG. 8 of the subject application, a 3D pipeline may include a bilinear interpolator, designated 820 and 830, such as one having a 6-bit interpolation phase value, as shown by FIG. 10 of the subject application. In such an embodiment, although the invention is not limited in scope in this respect, motion displacement and motion fractional offset may be calculated from motion vectors, decoded from the compressed bit stream as follows:

$$\begin{cases} D_x^Y = V_x >> SubD_x \\ D_y^Y = V_y >> SubD_y \end{cases} \text{ and } \begin{cases} D_x^C = (V_x / 2) >> SubD_x \\ D_y^C = (V_y / 2) >> SubD_y \end{cases} \quad [5]$$

and

$$\begin{cases} F_x^Y = (V_x \,\&\, FMaskD_x) << SubR_x \\ F_y^Y = (V_y \,\&\, FMaskD_y) << SubR_y \end{cases} \text{ and } \begin{cases} F_x^C = ((V_x / 2) \,\&\, FMaskD_x) << SubR_x \\ F_y^C = ((V_y / 2) \,\&\, FMaskD_y) << SubR_y \end{cases} . \quad [6]$$

For these relationships and this embodiment, values for the subsampled displacement shifts $SubD_x$ and $SubD_y$, the subsampled fractional masks $FMaskD_x$, $FMaskD_y$, and the subsampled bilinear interpolation phase shifters $SubR_x$, $SubR_y$,based at least in part on the downsampling ratio, may be determined. These are provided in Table 1, below, for a system with 6-bit interpolation phase value range. It will be appreciated that the values for a system with a different interpolation precision may, likewise, be derived as desired. It will also be appreciated the corresponding interpolation parameters for a system with a different interpolation filter other than a bilinear interpolation filter may also be derived as desired.

Table 1: Variables that are used to set the bilinear interpolation parameters for downsampling.

	Downsampling Ratio			
	1:1	2:1	4:1	8:1
$SubD_x$ or $SubD_y$	1	2	3	4
$FMaskD_x$ or $FMaskD_y$	0x01	0x03	0x07	0x0F

$SubR_x$ or $SubR_y$	5	4	3	2

With the above equations defining the motion displacement and motion fractional values, the motion-compensated prediction may be described for this embodiment by the following bilinear interpolation relation:

$$\begin{aligned} P[q][p] = \{ & (0x40 - F_y) \cdot [(0x40 - F_x) \cdot R[n][m] + F_x \cdot R[n][m+1]] \\ & + F_y \cdot [(0x40 - F_x) \cdot R[n+1][m] + F_x \cdot R[n+1][m+1]] \} // 0x80 \\ = \{ & \\ & (0x40 - F_y) \cdot [(0x40 - F_x) \cdot R[q + D_y][p + D_x] + F_x \cdot R[q + D_y][p + D_x + 1]] \\ & + F_y \cdot [(0x40 - F_x) \cdot R[q + D_y + 1][p + D_x] + F_x \cdot R[q + D_y + 1][p + D_x + 1]] \\ & \} // 0x80. \end{aligned} \qquad [7]$$

where, the reference pixel location (m, n) is derived from the motion vector displacement term (D_x, D_y). (Specification, page 11, line 17 to page 14, line 8; Figures 8 and 10).

FIG. 9 of the subject application illustrates the spatial relationship of the four reference pixels contributing to the prediction of the reconstructed pixel for this embodiment. As previously indicated, one embodiment of a bilinear interpolator is illustrated in FIG. 10 of the subject application. In this embodiment, bilinear interpolator 1001 is formed by two linear interpolators, 1020 and 1030, that operate along one spatial direction followed by another linear interpolator, 1040, that operates along the orthogonal direction. The output signal from linear interpolator 1040 passes through a rounding and saturation unit 1050 that converts the output signal to a specified finite precision form, although the invention is not limited in scope in this respect, and this is just one example of a bilinear interpolator embodiment. Furthermore, the

invention is not limited in scope to employing bilinear interpolation. (Specification, page 14, lines 9-17; Figures 9 and 10).

For this embodiment, the above mentioned motion compensation operation may be implemented with a hardware motion compensation system, such as the one, 701, illustrated in FIG. 7 of the subject application, although, again, the invention is not limited in scope in this respect. Here, the operation is applied on a macroblock basis, however, as previously indicated, this is merely a feature of MPEG2 and alternative embodiments are possible. The operation of this particular embodiment shall now be described in detail. Command parser and address generator unit 810 receives motion compensation instructions for a given macroblock and generates destination addresses and transmits the destination addresses to correction data memory interface unit 815 and destination data memory interface unit 825. Correction data memory interface unit 815 uses this(these) destination address(es) to load correction data from a correction data buffer(not shown). Destination data memory interface unit 825 uses this(these) destination address(es) to send the final output data from the motion compensation engine to the destination buffer(not shown). Command parser and address generator unit 810 also generates a prediction address (or addresses) in the reference picture or image using information about the current macroblock and its motion vectors and sends this to reference data memory interface unit 835. The reference data memory interface unit uses this to load data from a forward reference buffer, or from a backward reference buffer, or from both a forward reference buffer and a backward reference buffer(not shown). (Specification, page 14, line 18 to page 15, line 13; Figures 7).

The command parser and address generator unit also generates subpixel fractional information to be applied to the bilinear interpolation units, 820 and 830. Of these two bilinear interpolation units, one performs forward prediction and one performs backward prediction.

Here, each bilinear interpolation unit uses the fractional information to interpolate data from the reference buffer. It is conceivable that these two bilinear interpolation units may be implemented as a single hardware unit. In the case of a single hardware bilinear interpolation unit is implemented, this bilinear interpolation unit may be used sequentially if forward and backward bi-directional prediction is desired. (Specification, page 15, line 14 to page 16, line 2; Figure 7).

The output signals from the forward bilinear interpolation unit and the backward bilinear interpolation unit are added together in combine predictions unit 850. The combine predictions unit performs proper scaling and saturation to the data, such as according to a compression standard, such as, for example, MPEG2. The output signal from the combine predictions unit is then sent to prediction correction unit 860 and the correction data are added to the motion prediction data and final output data, for this embodiment, are generated. The output data from the prediction corrections unit is then sent to memory by the destination data memory interface. (Specification, page 16, lines 3-9; Figure 7).

As illustrated in FIG. 8 of the subject application, the above mentioned embodiment of motion compensation implementation may be implemented using existing 3D rendering hardware that is currently a common feature in graphics controller hardware. The boxes in dotted lines map the motion compensation aspects of this embodiment just described into some 3D hardware units, for illustration purposes. Of course, other hardware mapping and hardware reusing are also possible and may now be implemented by one of ordinary skill in the art. In this particular embodiment, the reference buffers are mapped as texture buffers. Therefore, the texture memory and texture cache may be used to obtain the reference data load from memory. After that, the 3D texture pipeline that typically contains bilinear interpolators or even tri-linear interpolators may be used to perform the bilinear interpolation and prediction combination operations in motion compensation. Then, the 3D texture blend unit may be used to perform the

prediction correction operation. The 3D color and destination memory interface unit may be used to write the output signals of motion compensation to memory. (Specification, page 16, line 10 to page 17, line 3; Figure 8).

Several embodiments where MPEG2 coding has been employed shall be described. As previously explained, the invention is not limited in scope to these particular embodiments. Any one of a number of other video or image coding specifications and/or storage formats may be employed. Nonetheless, these embodiments are provided as examples of implementations of a method of performing video image decoding in accordance with the present invention. In this context, three main categories of MPEG2 coding types shall be described. One coding type comprises a frame image with frame prediction or frame motion compensation employed. In this context, the term frame image or frame type refers to a progressive sequence display of data signals for an image, such as is commonly employed on computer platforms having monitors. The term frame prediction or frame motion compensation refers to a particular format for the prediction error and for the motion vectors that have been coded or produced by an encoder. It is desirable to know the format in which this signal information is encoded in the bit stream in order to perform decoding to reconstruct the image that produced this signal information. Therefore, if frame prediction or frame decoding is employed, then the prediction error is stored in a frame format, analogous to the format employed for a frame image. A second coding type comprises a field image with field motion compensation or field prediction. The term field image or field type generally refers to a technique commonly employed for television sets or television set displays in which half of the image is displayed separately at a rate that allows the human eye to merge the images. In this format, field data lines, that is, lines of signal data from a field image, are stored in an interlaced format. Therefore, top field and bottom field lines are alternated or interlaced within a frame of signal data. The term field motion compensation or

field prediction refers to the format in which the prediction error and motion vectors are stored in which prediction may be predicated upon the so-called top fields or bottom fields independently. In a field encoded image, the top and bottom fields are each encoded as separate images, and then displayed in an interlaced format. The motion prediction data for the top and bottom fields in this case is based in part on recently decoded fields. A third MPEG2 coding type employed in this context comprises a frame image with field motion compensation or field prediction. In this format, both fields are encoded as a single image, but the motion compensation data for each of its two fields is based in part on previously decoded fields. In MPEG2, this third format has two variations. In one variation, such as illustrated in Figure 5 of the subject application, the luminance DCT data is encoded on a frame basis, while in the other variation, such as illustrated in Figure 6 of the subject application, the luminance DCT data is stored on a field basis. This coding type is between the two coding types mentioned above in that both formats may be interspersed on a macroblock basis. More specifically, on a macroblock basis, data signals may be stored as a frame image with either field or frame prediction. (Specification, page 17, line 4 to page 18, line 18; Figures 5 and 6).

Because these particular embodiments relate to a DCT domain downsampling implementation for MPEG2 coding types, downsampling and motion compensation that is applied to the vertical direction will be employed. The horizontal direction in a video frame is handled similarly for the MPEG2 coding types described above, and therefore, in this embodiment, the horizontal direction is handled in a similar manner as the approach described below for a frame image with frame prediction, although, in a particular implementation of video image decoding in accordance with the present invention, this aspect may vary. Further, the illustrations given herein illustrate the technique for luminance component only. Nevertheless, an extension of this technique, once described, to handle the chrominance component of MPEG

is within the ability of one of ordinary skill in the art. Further, in other applications with multiple components encoded in the bit stream, such as, but not limited to, RGB encoded JPEG images, the extension of the technique described herein to each of the components is within the ability of one of ordinary skill in the art. (Specification, page 19, lines 1-12).

FIG. 3 of the subject application is a schematic diagram of an embodiment of a method for performing video image decoding in accordance with the present invention in which a DCT image that complies with the MPEG2 specification is employed. In this particular embodiment, a frame image with frame motion compensation, as described above, is the MPEG2 coding type employed. FIG. 3 of the subject application illustrates two 8x8 luminance blocks in a macroblock in which downsampling in the DCT domain occurs. Column 1 illustrates spatial positioning for data lines of the two blocks prior to downsampling. Column 2 illustrates spatial locations for the data lines after downsampling. Therefore, column 2 illustrates the effect on the data positioning of downsampling in the DCT domain and then performing the inverse DCT. Likewise, as FIG. 3 of the subject application illustrates, column 3 illustrates downsampling for a ratio of 4:1, as opposed to a ratio of 2:1 for column 2. As shown in FIG. 3 of the subject application, the downsampled lines are uniformly distributed in space after downsampling and inverse transforming. This would occur in this embodiment in a similar way for the downsampled pixels in the horizontal direction. Therefore, the downsampled frame image and frame motion vectors may be in a manner similar to the approach applied to the original image. The result of downsampling is to convert the 16x16 macroblocks and the 8x8 blocks they contain to smaller blocks. For example, after 2:1 horizontal subsampling and 4:1 vertical subsampling, each 8x8 block is decoded into a 4x2 block, and each 16x16 macroblock is decoded into an 8x4 macroblock. Thus, motion compensation for any given downsampled block, such as blocks with size 4x4, 4x2, 2x4, 2x2, or 1x1 in this embodiment, may be directly conducted on the

downsampled references using scaled motion vectors employing, in this particular embodiment, the technique described previously. Therefore, although the invention is not limited in scope in this respect, the previously described motion compensation hardware may be efficiently employed to perform this signal processing operation. (Specification, page 19, line 13 to page 20, line 15; Figure 3).

FIG. 4 of the subject application is a diagram illustrating an embodiment of a method for performing video image decoding in accordance with the present invention in which another MPEG2 coding type is employed. In this particular embodiment, a field image with field motion compensation is employed, as described above. Considering the nature of the two temporally separated fields for one frame, field based downsampling may introduce spatial aliasing and/or a non-uniform positioning of the lines from the two fields. The non-uniform positioning that may result is illustrated in FIG. 4 of the subject application in which, again, downsampling has been applied, and then the inverse DCT, to illustrate the effect on this coding type. However, the non-uniform line spacing does not affect motion vectors. Likewise, adjustments to the line positioning illustrated in FIG. 4 of the subject application, such as for the prediction error, may be accomplished using interpolation techniques such as bilinear interpolation. Again, the 3D hardware pipeline previously described may be employed to implement these interpolations. Therefore, in this particular embodiment, motion compensation, as well as the spatial positioning of the downsampled blocks, should include the exact line positioning for each field. (Specification, page 20, line 16 to page 21, line 10; Figure 4)

In another embodiment, instead of employing the approach illustrated in FIG. 4 of the subject application, which produced non-uniform vertical line spacing, selected taps may be employed for the top field and bottom field lines to produce a downsampled image that is uniformly spaced in the vertical direction. For example, although the invention is not limited in

scope in this respect, two spatial filters, one respectively for each of the bottom and top fields, may be employed. In addition, a similar approach may alternatively be employed in the frequency domain, such as the DCT domain. Where it is employed in the frequency domain, the transformed data signals may be phase shifted, rather than spatially shifted. The relation of a spatial shift and its corresponding transform domain operation may be derived using convolution property of the particular transform. (Specification, page 21, line 11 to page 22, line 1; Figure 4).

FIGs. 5 and 6 of the subject application each illustrate portions of embodiments of a method for performing video image decoding in accordance with the present invention for an MPEG2 coding type described as a frame image with field motion compensation. FIG. 5 of the subject application illustrates application of a portion of an embodiment of a method for performing video decoding in accordance with the invention to a macroblock stored in this format as a frame type with field prediction and frame DCT. In contrast, FIG. 6 of the subject application illustrates application of a portion of an embodiment of a method for performing video decoding in accordance with the present invention to a macroblock stored in this format as a frame type with field prediction and field DCT. It may be convenient to convert the image data and prediction or motion compensation data to one format, either frame or field. Likewise, conversion to a frame format may generally involve temporal filtering, which might involve a modification of the previously illustrated 3D pipeline hardware. However, of course, the invention is not limited in scope in this respect and this approach may be employed with a hardware pipeline, for example, that includes this feature. In this particular embodiment, however, operations are performed to place the frame data in a field format, and to place the frame motion compensation data into a field motion compensation format. Each field is then processed separately in the spatial domain to accomplish motion compensation, in this particular embodiment. (Specification, page 22, lines 2-17; Figures 5 and 6).

One modification, then, for this particular embodiment is to convert a frame downscaled macroblock into a field downscaled macroblock. In this particular embodiment, as illustrated in FIG. 5 of the subject application, this is accomplished by reconstruction of the blocks in a macroblock at full vertical resolution in the spatial domain by inverse transformation from the DCT domain, interlacing the block into two fields and downscaling it vertically in the spatial domain. Therefore, for this embodiment, the vertical downscaling is effectively moved to after performing the inverse DCT, as illustrated in FIG. 2 of the subject application. Likewise, motion compensation is performed on each field separately, as mentioned above. If the motion compensation were frame based, then, in this embodiment, the prediction error could be converted to field based using the technique illustrated. To convert frame motion vectors to field based, the frame motion vector may be employed for each of the top and bottom field motion vectors. A difference between the embodiments illustrated in FIG. 5 and FIG. 6 of the subject application is whether the macroblock is stored as a frame macroblock or a field macroblock. As previously discussed, if it is stored as a frame macroblock, then interleaving is performed as illustrated in FIG. 5 of the subject application. In contrast, as illustrated in FIG. 6 of the subject application, if the macroblock is stored as a field macroblock, then interleaving is performed, as illustrated, and the data lines may be processed as previously described for an interleaved field format. (Specification, page 22, line 18 to page 23, line 14; Figures 2, 5, and 6).

An aspect of an embodiment in accordance with the invention is the downscaling of a video image in the frequency domain, such as an MPEG2 image in the DCT domain, although the invention is not limited in scope in this respect. This may be discussed by referring to one-dimensional (1D) signals. The results for 2D signals would be an extension of this approach due to the separability of operations. Likewise, the case of 2:1 downscaling will be discussed as representative of other downscaling ratios. In general, implementing downscaling in the

frequency domain is well-known and there are many well-known ways to accomplish it. The invention is not restricted in scope to a particular approach and this discussion is provided as only one example. (Specification, page 23, line 15 to page 24, line 4).

The filtering of finite digital signals in the sample domain is performed using convolution. A well-known circular convolution may be obtained, for example, by a periodic extension of the signal and filter. It may be efficiently performed in the discrete Fourier transform (DFT) domain by simple multiplication of the discrete Fourier transforms of the signal and filter and then applying the inverse DFT to the result. For the DCT, a convolution may be applied that is related to, but different from the DFT convolution. This is described, for example, in "Symmetric Convolution and the Discrete Sine and Cosine Transforms," by S. Martucci, IEEE Transactions on Signal Processing, Vol. 42, No. 5, May 1994, and includes a symmetric extension of the signal and filter, linear convolution, and applying a window to the result. For example, assuming that the signal is represented as s(n), n=0,...,N-1, and its corresponding transform (DCT domain) coefficients is represented as S(u), u=0,..., N-1, and the filter is represented as h(m), m=0,...,M-1, then the DCT may be represented in matrix form as S=C*s, with s, S being column vector form of the signal and its DCT coefficients and C being the DCT matrix, as follows:

$$C_{u,n} = (2/N)^{1/2}\, k(u) \cos[\pi(u(2n+1)/2N)], \qquad \text{where } u,n = 0,...,N-1 \quad [8]$$

where

$$k(u) =$$

- $1/\sqrt{2}$, where $u = 0$ [9]
- 1, $u = 1,...,N-1$ (Specification, page 24, line 5 to page 25, line 5).

Assume a symmetric low pass even length filter h(m) with filter length M, where M=2*N, the DCT coefficients H(u) for the filter may be obtained by applying the convolutional form described above to the right half of the filter, which is equivalent to multiplication of the right half coefficients by the transform matrix:

$$D_{u,m} = 2\cos[\pi u(2m+1)/2n], \qquad \text{where } u,m = 1,...,N-1 \qquad [10]$$

The filtering is then performed by element-by-element multiplication of the signal DCT coefficients and the filter DCT coefficients and taking the appropriate inverse DCT transform of the DCT-domain multiplication results:

$$Y(u) = S(u) * H(u), \qquad \text{where } u = 0,....,N-1 \qquad [11]$$

Not only filtering, but also downsampling, may be performed in the DCT domain. For downsampling by two, the result of the element-by-element multiplication is folded across the middle half point and subtracted and after that scaled by $1/\sqrt{2}$. Mathematically, this is illustrated as:

$$[Y(u) - Y(N-u)]/\sqrt{2}, \qquad \text{where } u = 0,...(N/2)-1 \qquad [12]$$

(Specification, page 25, line 6 to page 26, line 2).

The decimated signal is then obtained by applying the inverse DCT transform of the length N/2. There are several special cases that might be usefully applied in this embodiment, although the invention is not limited in scope in this respect. For example, a brickwall filter with coefficients [11110000] in the DCT domain may be implemented that can further simplify the DCT domain downsampling by two operation. Specifically, the special filter shape avoids folding and addition. Another filter with coefficients [1 1 1 1 0.5 0 0 0] provides a transform

function of an antialising filter for the downsampling by two operation. Other filters may also be employed, of course. (Specification, page 26, lines 3-9).

Likewise, it will be appreciated that in this particular embodiment, a low pass, linear interpolation filter has been implemented to perform the downsampling; nonetheless, the invention is not limited in scope in this respect. For example, linear filters other than low pass filters or, alternatively, non-linear filters, such as, for example, a median filter, an adaptive edge-enhancement filter may be employed. It will, of course, be appreciated that some linear filters may effectively be implemented using motion compensation hardware and bilinear interpolation, although the invention is not limited in scope in this respect. (Specification, page 26, lines 10-16).

Filtering may also be applied after motion compensation or downsampling. More specifically, variations in clarity of the resulting images may become apparent to the human eye, particularly as the images are viewed in sequence. In some embodiments, it may be desirable to smooth these variations or, alternatively, enhance the images having less clarity. Therefore, any one of a number of filters, linear or non-linear, may be applied. For example, an edge enhancement image may be applied, although the invention is not limited in scope in this respect. Again, it will be appreciated that some linear filters may be effectively implemented using a 3D hardware pipeline and bilinear interpolation. (Specification, page 26, line 17 to page 27, line 6).

Of course, as previously indicated, the invention is not restricted in scope to the embodiments previously described. For example, in an alternative embodiment, where a 3D hardware pipeline is employed to implement a bilinear interpolation operation, a 3x3, 4x4, or greater interpolation operation may be implemented in place of a 2x2 bilinear interpolation operation. Likewise, in another alternative embodiment, as greater computational resources are demanded by the decoder in order to keep up with the video bit stream being provided or

received, the decoder may be adapted to downsample at higher ratios in order to allow graceful degradation in the quality of the images provided. Likewise, the decoder may be adapted to perform the reverse as well. (Specification, page 27, lines 7-14).

In another embodiment, instead of downsampling all video images, the decoder may be adapted to downsample only some of the video images. For example, specific images may be selected for downsampling, such as by transmitting a signal indication, or the decoder may be adapted to downsample a subset of the received video images based at least in part on a predetermined criteria, such as, as one example, decoding I and P frames at full resolution while subsampling B frames. Therefore, any one of a number of approaches may be employed and the invention is not restricted in scope to any particular approach. (Specification, page 27, line 15 to page 28, line 3).

Another aspect of an embodiment in accordance with the invention is the display of the decoded video images that are downsampled in the frequency domain, such as an MPEG2 image in the DCT domain, although the invention is not limited in scope in this respect. In this particular embodiment, the video decoder subsystem discussed above is coupled to a video display subsystem, as illustrated in FIG. 10 of the subject application. Both the video decoder subsystem and the video display subsystem may be coupled with the memory subsystem, where decoded video images may reside. As illustrated in FIG. 10 of the subject application, in the memory subsystem, the decoded video images are labeled as video buffer 1, video buffer 2 and so on. The number n of decoded video images may be chosen according to the video decoder and video display subsystems. In such an embodiment, besides typical information, such as the decoded image size (X, Y), the video decoder subsystem may be coupled with the video display subsystem with additional signals, such as the Picture Type (PICT) and the vertical subsampling factor (VSFF), that relate to the transform-domain downsampling operation. Signals such as

PICT and VSFF may be used to adjust the video display subsystem to properly display the decoded video images that are downsampled in the transform domain using an embodiment in accordance with the invention. (Specification, page 28, lines 4-17; Figure 10).

The video display subsystem handles displaying the decoded video images on the screen. The size of the desired display video window may not be the same as the source video image. In this case, the source video may be scaled up or down to match the display window size, corresponding to the process of interpolation and decimation, respectively. Quality scaling involves proper filtering of the source video data to reduce aliasing artifacts. In one approach, a finite impulse response (FIR) filter, where only finite number of input pixels contributes to a particular output pixel, is an example of a scaling filter implemented in the video display subsystem. A filter for spatial scaling of a video signal is normally a 2-dimensional (2D) function. In practice, separable filters may be used to reduce the hardware complexity and cost. In other words, the scaling of a video signal is applied to the vertical and horizontal directions independently. In the following, the vertical scaling operation is addressed since it is relevant to the uniform and non-uniform field scan line distribution that the proposed video decoder generates. (Specification, page 28, line 18 to page 29, line 11).

For a given source size N_{src} and a destination size N_{dest}, the forward scaling factor (in contrary to the backward scaling factor that we will define later) is defined as the ratio of the source size over destination size:

$$S_f = \frac{N_{src}}{N_{dest}}. \qquad [13]$$

Denoting the source sampling step as unity, we can define a DDA (Digital Differential Analyzer) value for a given output line as the relative position to the source line vertical positions. Normally, a DDA accumulator contains a fixed-point value. The integer portion of the DDA

value, denoted by int(DDA), indicates the closest source line number, while the fractional portion of the DDA value, denoted by fract(DDA), corresponds to the relative distance from that source line. The initial phase of a scaling operation is defined as the initial value of the DDA accumulator (DDA_0 = DDA(0)) that is associated with the first output line from the scaling filter. Then the sample position of a succeeding output line may be described by the DDA value accumulated by the scaling factor.

$$DDA(n) = DDA(n-1) + S_f, \quad \text{for } n = 1, N_{dest} - 1, \qquad [14]$$

where n is the index to the output video lines. (Specification, page 29, line 12 to page 30, line 7).

For a source video image that is created by the above mentioned video decoder subsystem and is in a frame type with the transform domain downsampling as illustrated in FIG. 3 of the subject application, its display is similar to the non-downsampled video image, although the scaling factor is different. (Specification, page 30, lines 8-10; Figure 3).

For a source video image that is created by the above mentioned video decoder subsystem and is in a field type with the transform domain downsampling but with uniformly distributed scan lines as illustrated in FIG. 12 of the subject application, again, its display method is similar to the non-downsampled field video image, although the scaling factor is different. However, for a source video image that is created by the above mentioned video decoder subsystem and is in a field type with the transform domain downsampling but with non-uniformly distributed scan lines, as illustrated in FIG. 13 and FIG. 14 of the subject application, the conventional field video display method cannot be applied to this kind of video images. Instead, proper vertical position adjustment is employed to display the top and bottom fields of the transform-domain downsampled video images correctly. (Specification, page 30, line 11 to page 31, line 1; Figures

12-14).

Let the distance between two adjacent lines in a field be 1 unit. As illustrated in FIG. 12 of the subject application, for the non-downsampled field-type video image, the first line in the bottom field (line 1) is 0.5 unit below the first line in the top field (line 0). This is also true for the subsequent lines in the top and bottom fields. The results of a DDA-based vertical scaling operation for uniformly-positioned interlaced video source are illustrated in FIG. 15 of the subject application. The example shows the upscaling factor of 3:8. FIG. 15 (a) of the subject application is the case of scaling from the top field with an initial phase of DDA[0] = 0.0, and FIG. 15(b) of the subject application is the case of scaling from the bottom field with an initial phase of DDA[0] = -0.5. (Specification, page 31, lines 2-8; Figures 12, 15, 15(a), and 15(b)).

When vertical downsampling by two is performed in the transform domain, the first line in bottom field (line 1) is 0.25 units below the first line in the top field (line 0) as illustrated in FIG. 13 of the subject application. FIG. 16 of the subject application illustrates the results of a DDA-based vertical scaling operation for non-uniformly-positioned interlaced video source. The example shows the upscaling factor of 3:8. FIG. 16 (a) of the subject application is the case of scaling from the top field with an initial phase of DDA[0] = 0.0, and FIG. 16 (b) of the subject application is the case of scaling from the bottom field with an initial phase of DDA[0] = -0.25. Similarly, FIG. 14 of the subject application illustrates that the first line in the bottom field is 0.125 units below the fist line in the top field when a vertical downsampling by four is performed in the transform domain. (Specification, page 31, line 9-16; Figures 13, 14, 16, 16(a), and 16(b)).

VI. ISSUES PRESENTED

A. Whether claims 1-7, 9, 11-12, 16-19, 21-24, 28-30, and 32-34 are patentable under 35 USC § 103 over the combination of Vetro et al. in view of Ng and Bose et al.

B. Whether claims 14 and 26 are patentable under 35 USC § 103 over the combination of Vetro et al., Bose et al., Ng, and Dugad et al.

C. Whether claims 8 and 31 are patentable under 35 USC § 103 over the combination of Vetro et al. and Ng, in further view of Kim et al.

D. Whether claims 15 and 27 are patentable under 35 USC § 103 over the combination of Vetro et al., Ng, Bose et al., and Dugad et al., in further view of Rosman et al.

VII. GROUPING OF CLAIMS

For each ground of rejection contested by Appellants in this appeal, the groupings of claims are as follows:

For purposes of the Examiner's rejection of claims 1-7, 9, 11-12, 16-19, 21-24, 28-30, and 32-34 under 35 USC § 103 as being obvious over the combination of Vetro et al. in view of Ng and Bose et al., claims 1-7, 9, 11-12, 16-19, 21-24, 28-30, and 32-34 stand or fall together as Group I.

For purposes of the Examiner's rejection of claims 14 and 26 under 35 USC § 103 as being obvious over the combination of Vetro et al., Bose et al., Ng, and Dugad et al., claims 14 and 26 stand or fall together as Group II.

For purposes of the Examiner's rejection of claims 8 and 31 under 35 USC § 103 as being obvious over the combination of Vetro et al. and Ng, in further view of Kim et al., claims 8 and 31 stand or fall together as Group III.

For purposes of the Examiner's rejection of claims 15 and 27 under 35 USC § 103 as being obvious over the combination of Vetro et al., Ng, Bose et al., and Dugad et al., in further view of Rosman et al., claims 15 and 27 stand or fall together as Group IV.

Reasons for the separate patentability of Claim Groups I-IV are presented in the Argument section pursuant to 37 C.F.R. § 1.192(c)(5).

VIII. ARGUMENT

A. THE COMBINATION OF VETRO ET AL., NG, AND BOSE ET AL. DOES NOT TEACH OR SUGGEST CLAIMS 1-7, 9, 11-12, 16-19, 21-24, 28-30, and 32-34

As the Honorable Board is well aware, in order to establish a *prima facie* case of obviousness:

> First, there must be some suggestion or motivation, either in the references themselves or in the knowledge generally available to one of ordinary skill in the art, to modify the reference or to combine reference teachings. Second, there must be a reasonable expectation of success. Finally, the prior art reference (or references when combined) must teach or suggest all the claim limitations. . . **The teaching or suggestion to make the claimed combination and the reasonable expectation of success must both be found in the prior art, not in applicant's disclosure** [emphasis added]. *In re Vaech*, 947 F.2d 488, 20 USPQ2d 1438 (Fed. Cir. 1991). *Manual of Patent Examining Procedure* (MPEP), 8th Edition, August 2001, § 2143.

In pertinent part, Vetro et al. teaches processing schemes that involve down conversion and motion compensation. **Significantly, as is acknowledged by the Examiner at pages 3 and 5 of the Final Office Action, Vetro et al. does not disclose "performing motion compensation for the downsampled image in the spatial domain, the performing of the motion compensation comprising scaling a motion vector in accordance with a downsampling ratio, the motion vector specifying relative distance of reference data from a macroblock,"** as is required in Appellants' independent claims. (See, independent claims 1, 16, 28, and 32).[1]

In pertinent part, Ng discloses decimating 8 by 8 blocks of data from VRAM by a decimator down to 4 by 4 blocks, and these 4 by 4 blocks are supplied to an adder in accordance with the data format of inverse transform image data applied to the adder from another decimator. (See, Ng, column 6, line 65 to column 6, line 7; See also, Ng's Figure 5).

Also in Ng, a motion compensated predictor receives "motion vectors" (as the term "motion vectors" is defined and used in Ng) and accesses blocks of pixel data **at addresses in VRAM identified by such "motion vectors."** See, e.g., Ng, column 3, line 64 to column 4, line 18, column 5, lines 38 to 64). As defined and used in Ng:

> Motion vectors . . . are codewords which identify 8 by 8 blocks of pixels in frames from which predicted frames are generated, which blocks most closely match the block currently being processed in the frame currently being encoded. (Ng, column 4, lines 34-39).

[1] At page 7 of the Final Office Action, the Examiner asserts that Vetro et al. inherently discloses "the DCT image is stored as complying with an MPEG specification" because "the Vetro method is to solve the drift and block artifact problems of MPEG-2." Contrary to the Examiner's assertion, Vetro et al.'s Abstract does not teach that Vetro et al. inherently discloses these features, and Appellants respectfully submit that Vetro et al. nowhere teaches that the DCT image in Vetro et al. is stored in compliance with an MPEG specification. During the prosecution of the subject application, Appellants have repeatedly traversed this assertion by the Examiner, and requested pursuant to MPEP § 2112, that the Examiner either withdraw this assertion, cite a specific passage in Vetro et al. that teaches that the DCT image in Vetro et al. is stored in compliance with an MPEG specification, or cite a specific passage in Vetro et

In making this final rejection, the Examiner asserts, "It would have been obvious . . . to include the scheme of Ng [or Bose al.] in the method of Vetro in order to improve the image recognition accuracy . . . Doing so would convert the format of the motion vector so as to improve accuracy of image reconstruction so that the quality of the method is improved."[2] Final Office Action, pages 6 and 9. Appellants respectfully traverse this assertion by the Examiner.

Contrary to the Examiner's assertions, the "motion vectors" disclosed in Vetro et al. are vastly different in purpose, effect, result, and operation from the "motion vectors" disclosed in Ng. (Vetro et al., Section 4.3, page 11, Bose et al., col. 17, lines 4-25, and Ng, col. 4, lines 34-39). Also contrary to the Examiner's assertion, there is no motivation or suggestion in any of the this prior art to selectively combine the teachings of Vetro et al., Bose et al., and Ng in the manner contemplated by the Examiner. Additionally, given the stark differences between the "motion vectors" disclosed in Vetro et al., Bose et al., and Ng, none of this prior art, whether taken singly or in any combination, can be said to suggest to those skilled in the art either the desirability of the selective combination of teachings of Vetro et al. and Ng proffered by the Examiner, or a reasonable likelihood of success of this selective combination.

Significantly, the Examiner acknowledges that neither Vetro et al. nor Ng disclose "the motion vector specifying relative distance of reference data from a macroblock." Final Office Action, page 6. However, the Examiner asserts, "the definition of the motion vector in MEPQ [sic, presumably 'MPEG'] is the relative distance from a reference macroblock to a predicted

---

al. that indicates that this is **necessary** to Vetro et al.'s disclosed techniques. To date, the Examiner has failed to provide a response to these requests.

[2] Contrary to the Examiner's assertion at page 6 of the Final Office Action, the portions of Bose et al. and Vetro et al. do not support the Examiner's assertion that ". . . Doing so would convert the format of the motion vector so as to improve accuracy of image reconstruction so that the quality of the method is improved." Final Office Action, page 6. During the prosecution of the subject application, Appellants repeatedly requested, pursuant to MPEP § 2144.03, that the Examiner supply a specific citation to a prior art reference to support the Examiner's assertion, or withdraw both the Examiner's assertion and the rejections based thereon. To date, the Examiner has yet to provide Appellants with a response to these requests.

macroblock. Thus, the claim language is inherent in the definition of the motion vector." Final Office Action, page 6.[3] However, the Examiner has provided no specific prior art citation to support this assertion. During the prosecution of the subject application, Appellants repeatedly requested, pursuant to MPEP § 2144.03, that the Examiner supply a specific citation to a prior art reference to support the Examiner's assertion, or withdraw the Examiner's assertion. To date, contrary to the provisions of the MPEP, the Examiner failed to supply any such evidence, but nevertheless, maintains his factually unsupported claim rejections!

Additionally, as stated above, the respective definitions and uses of the "motion vectors" disclosed in Vetro et al. and Ng are irreconcilably different from each other. Accordingly, as a matter of logic, the Examiner's assertion that the definition of the claim language "motion vector" recited in Applicant's claims is "inherently" disclosed in both Vetro et al. and Ng clearly is in error.

Concerning Bose et al., the Examiner asserts, "The limitation [that is acknowledged by the Examiner to be missing from Vetro et al. and Ng] is well known in the art. Bose, in an analogous environment, explicitly teaches [this] limitation . . ." Final Office Action, page 6. Even assuming, for the sake of argument, that Bose et al. discloses a motion vector that specifies relative distance of reference data from the processed macroblock, none of the prior art relied upon the Examiner suggests the selective combination of teachings of Vetro et al., Ng, and Bose et al. relied upon by the Examiner. Indeed, even as characterized by the Examiner, the respective definitions and uses of the "motion vectors" disclosed in Vetro et al., Ng, and Bose et al. are mutually different, and are irreconcilably inconsistent with each other. See, Final Office Action, pages 2 and 6. No guidance is supplied by any of the prior art relied upon by the Examiner that

[3] Apparently in response to Appellants' previous arguments during the prosecution of the subject application, the Examiner appears to have withdrawn the Examiner's previous assertions that Ng discloses an MPEG or MPEG-like

would resolve their mutually contradictory teachings so as to suggest the selective combination of teachings proffered by the Examiner. Additionally, given the radical differences between the "motion vectors" disclosed in Vetro et al., Ng, and Bose et al., none of this prior art can be said to suggest to those skilled in the art a reasonable likelihood of success of this selective combination proffered by the Examiner.

In a vain effort to try to overcome these deficiencies of Vetro et al., Ng, and Bose et al., the Examiner asserts:

> In the instant case, the motion vectors in Ng are codewords which identify 8 x 8 block [sic] of pixels as applicant pointed out, while the motion vectors in Vetro is [sic] vectors (or codewords) used to specify the neighborhood of blocks . . . Thus, the motion vectors in Ng and Vetro have no significant difference to prevent the combination of references. Final Office Action, page 2.

However, contrary to the Examiner's assertion, even a casual review of the definitions of the term "motion vector" given in Vetro et al., Ng, and Bose et al. (and, for that matter, as characterized by the Examiner) reveals that these definitions cannot properly be characterized as lacking "significant differences" that would prohibit the selective, mosaic, combination proffered by the Examiner. This clearly evidences the fact that the Examiner's selective combination of Vetro et al., Ng, and Bose et al. is based upon improper hindsight!

Thereafter, the Examiner asserts "The motivation to combine Bose [with Vetro et al. and Ng] is explicitly disclosed in Bose and Vetro (Bose, col. 3 to col. 4, Vetro, abstract)." Final Office Action, page 3. However, even a casual reading of these portions of Bose et al. and Vetro

---

system. However, the Examiner has yet to provide any evidence to support the Examiner's suggestion in the Final Office Action that the MPEG definition of "macroblock" is inherently disclosed in Ng.

et al. reveals that these portions of Bose et al. and Vetro et al. do not explicitly provide such motivation, but instead, clearly evidences that the Examiner has relied upon improper hindsight. Furthermore, even assuming, for the sake of argument, that Vetro et al., Ng, and Bose et al. are in the same field and/or address the same problem to be solved, this would only evidence that they might arguably constitute mutually analogous art. Unless they also provide motivation for the selective combination proffered by the Examiner (which they clearly do not), the selective combination proffered by the Examiner cannot properly be used to reject the claims under 35 USC § 103.

Accordingly, the proffered selective combination of teachings of Vetro et al., Ng, and Bose et al. cannot render obvious claims 1-7, 9, 11-12, 16-19, 21-24, 28-30, and 32-34. Therefore, it is respectfully submitted that the Examiner's final rejection of claims 1-7, 9, 11-12, 16-19, 21-24, 28-30, and 32-34 under 35 USC § 103 as being rendered obvious by this selective combination is erroneous, and should be reversed.

B. THE COMBINATION OF VETRO ET AL., BOSE ET AL., NG, AND DUGAD ET AL. DOES NOT TEACH OR SUGGEST CLAIMS 14 AND 26.

The deficiencies of Vetro et al., Bose et al., and Ng vis-à-vis the independent claims of the subject application have been previously described in detail. It is not seen that Dugad et al. overcomes these deficiencies of Vetro et al., Ng, and Bose et al. so as to suggest, when taken in combination with Vetro et al., Ng, and Bose et al., Appellants' claimed invention.

Claims 14 and 26 depend directly from independent claims 1 and 16, respectively, and therefore, must be read as incorporating the limitations of claims 1 and 16, respectively. 35 USC § 112, fourth paragraph. Dugad et al. is cited by the Examiner as disclosing the feature of using a bilinear interpolation scheme for downsampling. Even assuming, for the sake of argument, that

Dugad et al. discloses this feature, Dugad et al. nowhere discloses or suggests the aforesaid combination of limitations of Appellants' independent claims 1 and 16 that are missing from Vetro et al., Ng, and Bose et al.

Furthermore, none of this prior art can be said to provide any guidance that would resolve the mutually contradictory teachings of Vetro et al., Ng, and Bose et al., so as to suggest the selective combination of teachings of Vetro et al., Ng, and Bose et al. proffered by the Examiner. Also, none of this prior art can be said to teach or suggest, whether taken alone or in any combination, a reasonable likelihood of success of the selective combination proffered by the Examiner.

Additionally, the Examiner's rejection of claims 14 and 26 is based on no less than four isolated prior art documents that contain no teaching or suggestion to selectively combine their respective teachings in the manner contemplated by the Examiner. Clearly, this evidences the fact that the Examiner has engaged in improper hindsight to reject claims 14 and 26!

Accordingly, it cannot be said that claims 14 and 26 are rendered obvious by the combination of Vetro et al., Ng, Bose et al., and Dugad et al. Therefore, it is respectfully submitted that the Examiner's final rejection of claims 14 and 26 under 35 USC § 103 as being rendered obvious by the combination of Vetro et al., Ng, Bose et al., and Dugad et al. is erroneous, and should be reversed.

C. THE COMBINATION OF VETRO ET AL., NG, AND KIM ET AL. DOES NOT RENDER OBVIOUS CLAIMS 8 AND 31.

Claims 8 and 31 depend indirectly from claim 1, and thus, must be read as incorporating the limitations of claim 1. 35 USC § 112, fourth paragraph. Kim et al. fails to cure the

deficiencies pointed out above regarding Vetro et al. and Ng vis-à-vis claim 1. Kim et al. is cited by the Examiner as disclosing the feature of displaying a downsampled spatial image such that the resulting non-uniform vertical spacing of data signal lines appear substantially uniform on a low resolution monitor screen. Even assuming, for the sake of argument, that Kim et al. discloses this feature, Kim et al. nowhere discloses or suggests the aforesaid combination of limitations of Appellants' independent claims 1 and 16 that are missing from Vetro et al. and Ng.

Furthermore, none of this prior art can be said to provide any guidance that would resolve the mutually contradictory teachings of Vetro et al. and Ng so as to suggest the selective combination of teachings of Vetro et al. and Ng proffered by the Examiner. Also, none of this prior art can be said to teach or suggest, whether taken alone or in any combination, a reasonable likelihood of success of the selective combination proffered by the Examiner.

Accordingly, it cannot be said that claims 8 and 31 are rendered obvious by the combination of Vetro et al., Ng, and Kim et al. Therefore, it is respectfully submitted that the Examiner's rejection of claims 8 and 31 under 35 USC § 103 as being rendered obvious by the combination of Vetro et al., Ng, and Kim et al. is erroneous, and should be reversed.

D. THE COMBINATION OF VETRO ET AL., NG, BOSE ET AL., DUGAD ET AL., AND ROSMAN ET AL. DOES NOT TEACH OR SUGGEST CLAIMS 15 AND 27.

Claims 15 and 27 depend indirectly from claims 1 and 16, respectively, and thus, must be read as incorporating the limitations of claims 1 and 16, respectively. 35 USC § 112, fourth paragraph. Rosman et al. fails to cure the deficiencies pointed out above regarding Vetro et al., Ng, Bose et al., and Dugad et al. vis-à-vis claims 1 and 16. Rosman et al. is cited by the Examiner as disclosing the feature of using a 3D pipeline to perform bilinear interpolation. Even assuming, for the sake of argument, that Rosman et al. discloses this feature, Rosman et al.

nowhere discloses or suggests the aforesaid combination of limitations of Appellants' independent claims 1 and 16 that are missing from Vetro et al., Ng, Bose et al., and Dugad et al.

Furthermore, none of this prior art can be said to provide any guidance that would resolve the mutually contradictory teachings of Vetro et al., Ng, and Bose et al., so as to suggest the selective combination of teachings of Vetro et al., Ng, and Bose et al. proffered by the Examiner. Also, none of this prior art can be said to teach or suggest, whether taken alone or in any combination, a reasonable likelihood of success of the selective combination proffered by the Examiner.

Additionally, **this rejection is based upon a combination of no less than five isolated prior art documents that contain no teaching or suggestion to combine their respective teachings in the manner contemplated by the Examiner**. Clearly, the Examiner has engaged in hindsight to try to pick through isolated prior art disclosures in a vain effort to reconstruct the limitations of Appellants' claims!

Accordingly, it cannot be said that claims 15 and 27 are rendered obvious by the combination Vetro et al., Ng, Bose et al, Dugad et al., and Rosman et al. Therefore, it is respectfully submitted that the Examiner's rejection of claims 15 and 27 under 35 USC § 103 as being rendered obvious by the combination of Vetro et al., Ng, Bose et al., Dugad et al., and Rosman et al. is erroneous, and should be reversed.

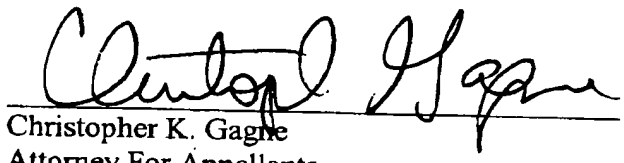
IX. CONCLUSION

For the foregoing reasons, Appellants respectfully submit that each and every one of the final rejections made by the Examiner in the Final Office Action is erroneous. Accordingly, Appellants respectfully request that the Honorable Board of Patent Appeals and Interferences reverse the Examiner and direct that all of the currently pending claims be allowed.

Please charge any shortages and credit any overcharges to Deposit Account number 02-2666.

Respectfully submitted,

Date: 28 May 2003



Christopher K. Gagne
Attorney For Appellants
Reg. No. 36,142

c/o Blakely, Sokoloff, Taylor, & Zafman, LLP
12400 Wilshire Boulevard, Seventh Floor
Los Angeles, CA 90025-1026
(503) 684-6200

APPENDIX A: CLAIMS ON APPEAL

1. A method of performing video image decoding comprising:

downsampling a compressed video image in the frequency domain;

inverse transforming the downsampled video image; and

performing motion compensation for the downsampled image in the spatial domain, the performing of the motion compensation comprising scaling a motion vector in accordance with a downsampling ratio, the motion vector specifying relative distance of reference data from a macroblock.

2. The method of claim 1, wherein the compressed video image in the frequency domain comprises a discrete cosine transform (DCT) image.

3. The method of claim 2, wherein the DCT image is stored as a DCT image that complies with an MPEG specification.

4. The method of claim 3, wherein the DCT image is stored as a frame type image.

5. The method of claim 4, wherein the motion compensation data signals are stored as frame prediction type motion compensation.

6. The method of claim 3, wherein the DCT image is stored as a field type image.

7. The method of claim 6, wherein the motion compensation data signals are stored as field prediction type motion compensation.

8. The method of claim 1, and further comprising displaying the downsampled spatial image so that resulting non-uniform vertical spacing of data signal lines in the downsampled spatial image appear substantially uniform on a computer monitor.

9. The method of claim 1, wherein the downsampling is performed using an integer ratio.

11. The method of claim 1, wherein motion vector scaling comprises implementing an interpolation operation.

12. The method of claim 11, wherein motion vector scaling comprises implementing a bilinear interpolation operation.

13. The method of claim 12, wherein the bilinear interpolation operation is implemented on 3D pipeline hardware.

14. The method of claim 1, wherein downsampling comprises implementing a linear filter as a bilinear interpolation operation.

15. The method of claim 14, wherein the bilinear interpolation operation is implemented on 3D pipeline hardware.

16. A method of performing video image decoding comprising:

inverse transforming a compressed video image;

downsampling the inverse transformed image in the spatial domain; and

performing motion compensation for the downsampled image in the spatial domain, the performing of the motion compensation comprising scaling a motion vector in accordance with a downsampling ratio, the motion vector specifying relative distance of reference data from a macroblock.

17. The method of claim 16, wherein the compressed video image comprises a discrete cosine transform (DCT) image.

18. The method of claim 17, wherein the DCT image is stored as a DCT image that complies with an MPEG specification.

19. The method of claim 18, wherein the DCT image comprises macroblocks stored as frame macroblocks and macroblocks stored as field macroblocks.

20. The method of claim 19, and further comprising: converting the frame macroblocks to field macroblocks prior to downsampling in the spatial domain.

21. The method of claim 19, wherein the motion compensation data signals are stored as field prediction type motion compensation.

23. The method of claim 16, wherein motion vector scaling comprises implementing an interpolation operation.

24. The method of claim 23, wherein motion vector scaling comprises implementing a bilinear interpolation operation.

25. The method of claim 24, wherein the bilinear interpolation operation is implemented on 3D pipeline hardware.

26. The method of claim 16, wherein downsampling comprises implementing a linear filter as a bilinear interpolation operation.

27. The method of claim 26, wherein the bilinear interpolation operation is implemented on 3D pipeline hardware.

28. An article comprising: a storage medium, having stored thereon instructions, that when executed by a platform, result in the following:

downsampling a compressed video image in the frequency domain;

inverse transforming the downsampled video image; and

performing motion compensation for the downsampled image in the spatial domain, the performing of the motion compensation comprising scaling a motion vector in accordance with a downsampling ratio, the motion vector specifying relative distance of reference data from a macroblock.

29. The article of claim 28, wherein the instructions, when executed further result in the compressed video image in the frequency domain comprising a discrete cosine transform (DCT) image.

30. The article of claim 29, wherein the instructions, when executed, further result in the DCT image being stored as a DCT image that complies with an MPEG specification.

31. The article of claim 28, wherein the instructions, when executed, further result in: displaying the downsampled spatial image so that resulting non-uniform vertical spacing of data signal lines in the downsampled spatial image appear substantially uniform on a computer monitor.

32. An article comprising: a storage medium, having stored thereon instructions, that when executed by a platform, result in the following:

inverse transforming a compressed video image;

downsampling the inverse transformed image in the spatial domain; and

performing motion compensation for the downsampled image in the spatial domain, the performing of the motion compensation comprising scaling a motion vector in accordance with a downsampling ratio, the motion vector specifying relative distance of reference data from a macroblock.

33. The article of claim 32, wherein the instructions, when executed further result in the compressed video image in the frequency domain comprising a discrete cosine transform (DCT) image.

34. The article of claim 33, wherein the instructions, when executed, further result in the DCT image being stored as a DCT image that complies with an MPEG specification.